Long-term tracking of social structure in groups of rats

Máté Nagy[1,2,3,4,5,6]◑*, Jacob D. Davidson[4,5,6]◑*, Gábor Vásárhelyi[1,3], Dániel Ábel[1], Enikő Kubinyi[7,8,9], Ahmed El Hady[4,5,6], Tamás Vicsek[1,3]

1 Department of Biological Physics, Eötvös Loránd University, Budapest, Hungary
2 MTA-ELTE 'Lendület' Collective Behaviour Research Group, Hungarian Academy of Sciences, Budapest, Hungary
3 MTA-ELTE Statistical and Biological Physics Research Group, Hungarian Academy of Sciences, Budapest, Hungary
4 Department of Collective Behaviour, Max Planck Institute of Animal Behavior, Konstanz, Germany
5 Department of Biology, University of Konstanz, Konstanz, Germany
6 Centre for the Advanced Study of Collective Behaviour, University of Konstanz, Konstanz, Germany
7 Department of Ethology, Eötvös Loránd University, Budapest, Hungary
8 ELTE NAP Canine Brain Research Group, Budapest, Hungary
9 MTA-ELTE Lendület 'Momentum' Companion Animal Research Group, Budapest, Hungary

◑These authors contributed equally to this work.
* nagymate@hal.elte.hu, jdavidson@ab.mpg.de

Abstract

Rodents serve as an important model for examining both individual and collective behavior. Dominance within rodent social structures can determine access to critical resources, such as food and mating opportunities. Yet, many aspects of the intricate interplay between individual behaviors and the resulting group social hierarchy, especially its evolution over time, remain unexplored. In this study, we utilized an automated tracking system that continuously monitored groups of rats for over 250 days to enable an in-depth analysis of individual behavior and the overarching group dynamic. We describe the evolution of social structures within a group and investigate how past behaviors influence the emergence of new social hierarchies when group composition changes. Notably, a comparison of behavior in groups to conventional individual and pairwise behavioral assays reveals that the latter are insufficient for capturing or predicting genuine behavioral outcomes in a collective context. These results emphasize the context-dependence of social behavior as an emergent property of interactions within a group and highlight the need to measure and quantify social behavior in more naturalistic environments.

Author summary

Despite the fact that rodents are a model system for studying both individual and social behavior, testing methods are typically limited to short-term and simple behavioral tests. As a result, it is uncertain how different behavioral patterns interact to create a group's overall social structure. Using long-term experiments of over 250 days, we employed automated tracking to study rats' social behavior. We describe distinct

aspects of a group's social structure by employing different quantitative metrics. These methods allowed us to compare group behavior with traditional individual and pairwise behavioral tests. The results indicate that simplified assays do not accurately describe or predict actual behavioral outcomes in a group setting, and emphasize the need to measure and quantify behavior in naturalistic settings with long-term tracking in order to understand group social structure and behavior. Our research establishes a framework for characterizing the social structure of rats and other social animals.

Introduction

Collective behavior emerges based on interactions between individuals in a group. This is observed at many different scales, from wound healing at the cellular level [1], to task allocation in social insects [2], group search behavior [3], and information exchange on human social networks [4]. Hierarchical structures are common in animal groups, for example, in the grooming relationships of chimpanzees [5], leadership and movement of pigeons [6], and reproduction in cichlid fish [7]. Social network structure influences decision-making [8, 9], and dominance position within a network can influence an individual's fitness [10].

With rats, previous work has shown that individuals within a group have a social status related to dominance and that aggression and avoidance behavior are key elements of social interactions [11–13]. However, it is not known how individual interactions lead to the overall social structure of the group and how social structures change over time [14]. Fortunately, new automated tracking methods enable long-term tracking of individuals within social groups, providing a quantitative description of behavior and interactions [15]. For example, recent work has analyzed the ontogeny of collective behavior in zebrafish [16], lifetime behavioral differences in honeybees [17], and how genetic relatedness corresponds to group social structure in mice [18].

The study of social behavior is particularly important in animal models utilized for the understanding and treatment of social-related neuropsychiatric disorders. Rodents such as mice and rats have been indispensable as model biological organisms [19–21], with particular relevance to clinical research due to their short life span and tolerance to laboratory environments. However, the laboratory environment typically restricts the development of complex social behaviors. Under ideal circumstances, laboratory rats would be housed in complex environments in large groups that mimic the conditions experienced by their wild ancestor, the Norway rat (*Rattus norvegicus*) [11, 22]. During experiments, however, rats are often kept in small cages, and then transferred to separate environments in order to examine social interactions in the form of simplified behavioral assays such as the 3-chambered social preference test [23]. It is therefore not surprising that the translational relevance of these tests is limited [24, 25]. Incorporating environmental and social complexity into experiments can increase the generalizability of conclusions drawn from laboratory studies [26, 27].

While it is acknowledged that long-term multi-modal characterizations are required to describe complex social behaviors, there are still a number of challenges [28]. An approach applicable to both lab and field conditions is to tabulate appropriate pairwise interaction information of animals in groups and use this to define measures of an individual's position in the social hierarchy. With mice, both aggressive and non-aggressive interactions have been used to define dominance hierarchies [29–35]. With primates, for example, pairwise "supplanting" interactions, such as when one individual displaces another from a food source, have been used to determine an individual's ranking [36–38].

There are multiple ways to calculate social rankings based on the results of pairwise interaction matrices. The Elo score, which was originally developed to rank chess

players and predict the outcome of future matches [39], is commonly used in animal behavior to calculate an individual's ranking using pairwise contest or interaction information [40–42]. If network "flow" is defined from winners to losers in an interaction network, the network measure of "flow hierarchy," which refers to how information flows through the network, can also be used to quantify hierarchical structure. The local reaching centrality (LRC) considers flow hierarchy in a network and quantifies a node's ability to efficiently reach other nodes in its immediate neighborhood [43]. Since "flow" occurs in the direction from winner to loser (or dominant to less-dominant), dominant individuals have higher local reaching centrality. Global reaching centrality (GRC) is calculated using the distribution of LRC scores, and this metric has been used to quantify the steepness of hierarchical structures in many different systems, including groups of horses [44], ant colonies [45], brain networks [46], industrial trade networks [47, 48], and scientific citation networks [49]. In animal behavior, in addition to the individual Elo scores, metrics employed to measure social dominance hierarchies also include directional consistency of contests, proportions of contests based on rank, transitivity, linearity of the hierarchy, and David's score or Elo score distributions across the group [33, 50–53].

Automated identification of approach-avoidance behavior has been used in previous work as a scalable method to characterize group hierarchy and leader-follower relationships [6]. Previous work with rodents, including rats [54] and mice [15], has also considered approach and avoidance events. An approach-avoidance event occurs when one individual approaches another, but the other individual moves away (either by retreating or escaping). Because it only uses location and trajectory data, the use of approach-avoidance interaction matrices with network-based algorithms to determine ranking within a hierarchy has the advantages of being scalable, easy to implement, and generalizable.

In this work, we develop and apply an open-source vision-based automated tracking and behavioral characterization system designed to analyze the social behavior of small animals like rats. We use this system to characterize group social network structure, individual behavior and social ranking within each group, and the stability of social networks and behavior over time. For an extended 36-week period, we tracked the social behavior of 28 rats divided into several groups, and calculated behavioral metrics and interactions to analyze both individual and group behavior. We examined how individual behavioral differences persist and combine to form new social structures when the composition of the group is altered. At first, the rats were divided into 4 groups of 7, and following this, we merged groups. For the final sequence of group experiments we then created 4 new groups of 7. Following the completion of the group experiments, we ran individual behavioral assays on each rat and compared the results to those of the group experiments. This combination of methods and experiments enabled us to (1) identify a wide range of individual locomotion and social behaviors, (2) investigate the formation and details of dominance hierarchies, (3) investigate the effect of group composition changes and the associated social stress on the behavior of rats living in groups in semi-natural environments, and (4) compare behavioral assay results to behavior observed in a group setting. Overall, our work demonstrates scalable methods for describing long-term changes in animal group social structure and emphasizes the need to use such methods to obtain a full picture of group social structure and interactions in natural or semi-natural environments.

Results

Long-term tracking and quantifying individual behavior

We tagged individuals with color markers and employed automated tracking to determine each rat's movement over time (Figure 1A-D). Over the course of the experiment, we performed manipulations to alter the group composition and the living area available for the group to use. We used two different strains of laboratory rats, denoted A and B (see Methods for details), with associated individual labels $a1$-$a14$ and $b1$-$b14$, respectively (see Figure 1E). We initially divided the rats into four groups of seven, with A rats in groups $A1$ and $A2$ and B rats in groups $B1$ and $B2$. The rats remained in these groups for the first observation period, which lasted a total of 21 weeks - we denote this time as Period 1. Following this, in Period 2, we merged groups $A1$-$A2$ and $B1$-$B2$ for three weeks and then merged all for three weeks by opening portals between their compartments. For the final series of group experiments in Period 3, rats from each original group were mixed together to create four new groups. Figure 1E shows the experimental structure and the associated measurement periods.

We measured body mass, counted wounds, and calculated automated trajectory-based behavioral metrics to quantify behavior over the duration of the experiment. We calculated and averaged each metric over successive time periods of 3 weeks and refer to each 3-week period as an "Experiment" (denoted as Exp), with associated numbers 1-12. We use the summary metrics to ask how behavior changes over time, how individuals differ, how groups differ, and how previous individual and group behavior predicts changes when new groups are formed.

To assess dominance-related interactions and social structure, we tabulated approach-avoidance contests between all pairs of rodents in each group. This automated method defines "contests" as when a pair of rats come close to each other: the "displacer", i.e. the dominant rat in a contest, subsequently stays in place or continues moving forward, while the other (the "displaced", i.e. non-dominant rat) move away [6]. This type of approach-avoidance interaction can also be dynamic, such as when one individual chases another. We use the matrix of approach-avoidance contest outcomes to calculate metrics that describe the dominance structure of each group and each individual's position in this structure.

Strain and group differences

We use automated measures of space use and pairwise contests, in addition to manual body mass and wound count measurements, to characterize individual and group behavior. We first examine general differences between the groups that were organized by strain during Period 1.

In the beginning, the rats were juveniles and were growing rapidly, as shown by the large increases in body mass during this time period. The A rats were, on average, larger than the B rats. Except for rats $a9$ and $a10$ in Group $A2$, there were few wounds seen on the rats in all groups during Period 1. We saw an increase in the number of wounds and the number of approach-avoidance contests in Periods 2 and 3 (Figure 2).

We next compare the average space use across rats in each of the groups. In Period 1, rats in strain A groups tended to be farther from the wall, visited more parts of the living compartment (higher territory size), and spent less time on top of the nestbox in comparison with B rat groups (t-test, $p<0.01$ comparing per-rat averages of strain A versus B rats for distance from wall, territory size, and top of nestbox; Fig S2). We see that strain differences in space use persisted during Exp 8 when groups of the same strain were mixed. When all rats were mixed into a single group during Exp 9, strain and individual differences persisted but decreased (see also Fig S3). Following the

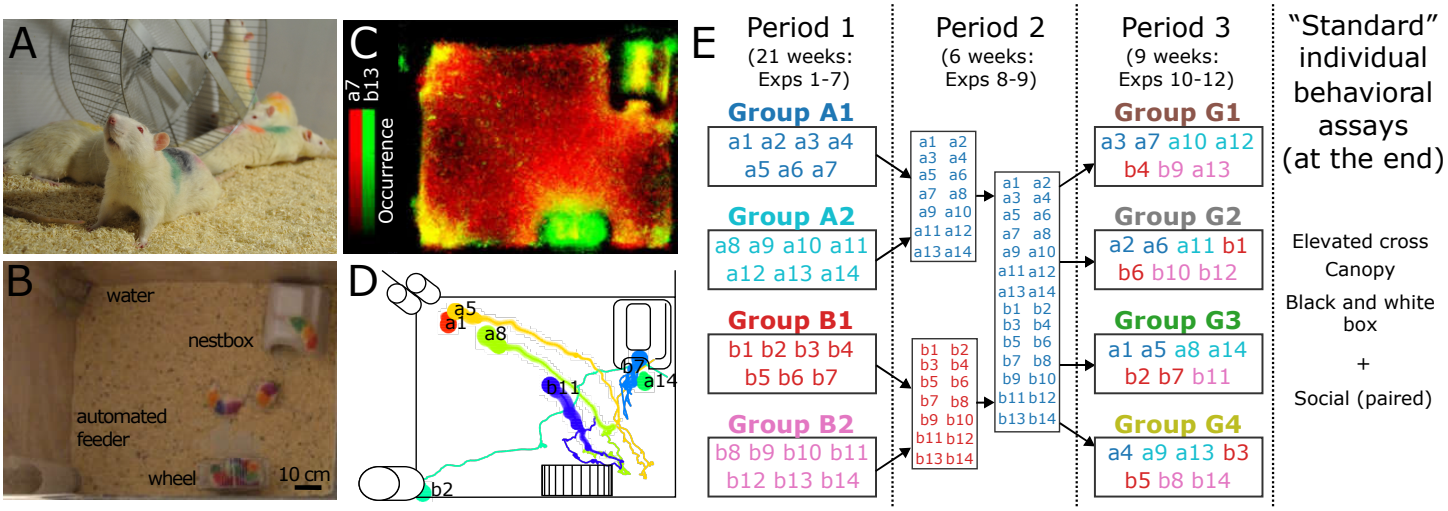

Fig 1. Experiment setup and timeline. (A) Photo of the rats with color-codes for individual identification and tracking. (B) Still image from the video that was used for tracking (from group $G1$, during Exp 10) taken by a light-sensitive camera at low lighting conditions. Image overlaid with labels indicating the important objects (water, nestbox, etc.). (C) Continuous tracking allowed for the reconstruction of each individual's space use. The heatmap shows the space use of two rats during a 3-week period at the beginning of Period 3. Areas used only by $a7$ are shown with red, only by $b13$ with green, and areas visited by both (e.g. at the water and the feeder) are shown with yellow. (D) Trajectories were used to identify dominance interactions in the form of approach-avoidance events, where one individual approaches another, but the other moves away (by backing up or fleeing). Shown is an example of trajectories from group $G3$ in Exp 10. Lines show locations for 60 seconds, with the semitransparent circles of increasing size showing the more recent positions. (E) Overview of experimental manipulations. We refer to each 3-week period as an "Experiment", and calculate behavioral metrics over these periods. Period 1 had rats in original strain-sorted groups $A1$-$A2$ (strain A), $B1$-$B2$ (strain B), for a total of 7 experiments. Each rat is labeled with lowercase letters a or b according to strain. Individual numbers within each group are sorted in ascending order according to body mass at the end of Period 1, i.e. $a7/a1$ were the largest/smallest in $A1$ during Exp 7, $a14/a8$ were the largest/smallest in $A2$, etc. In Period 2, the groups were mixed together by strain during Exp 8, and then all together for Exp 9. At the beginning of Period 3 (Exp 10), new groups were formed ($G1$-4). During Exps 11 and 12 in Period 3, the compartment area sizes were changed (see Methods and Fig S8). At the end of the experiments, individual behavior was assessed by traditional individual and pairwise assays.

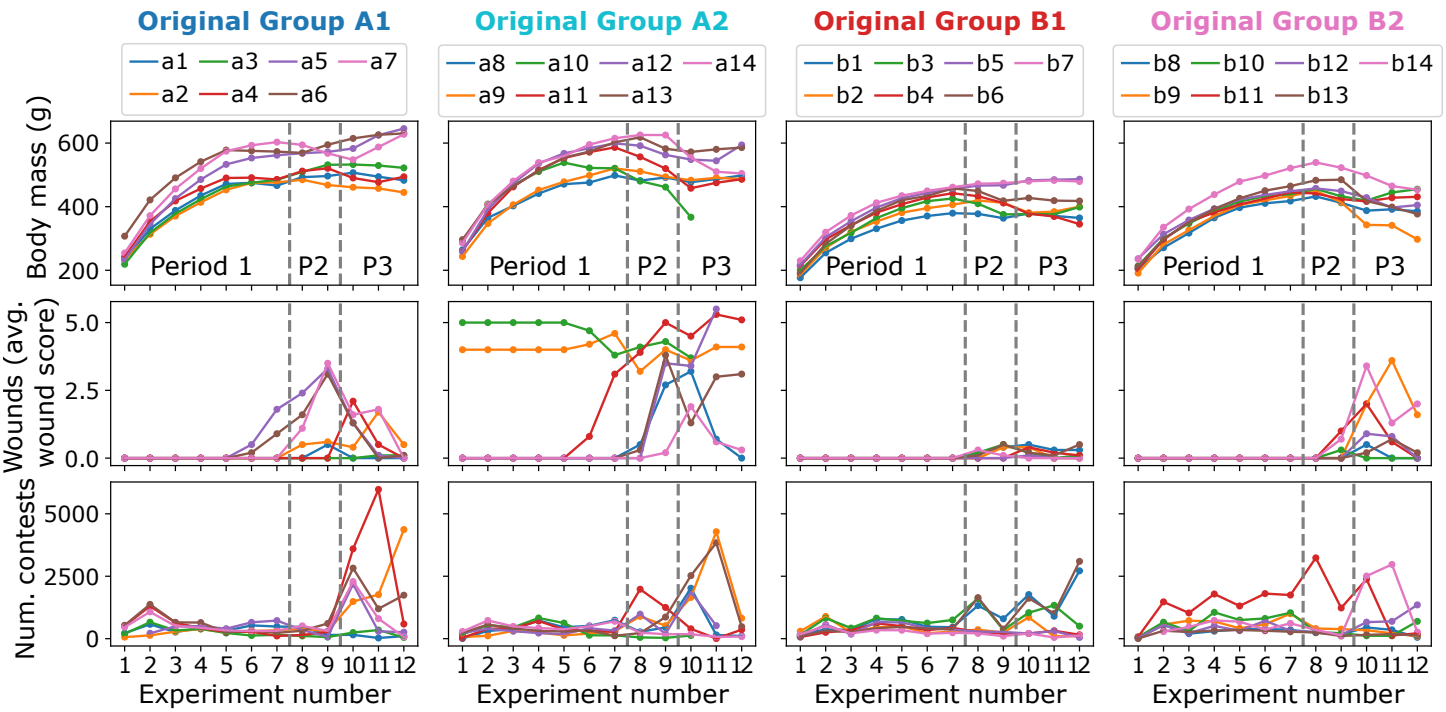

Fig 2. Body mass, wounds, and number of contests during the entire experiment. Individual rats are plotted according to Period 1 groups and shown with different colored lines. Num. contests are the mean number of pairwise contests one rat had with other rats in the same group. Wound scores were assigned according to the number of wounds counted, and values averaged over measurements during each Experiment (see Methods).

re-grouping in Period 3, the new groups (which were mixed with both A and B strain rats) differed in their space use tendencies, and the standard deviation of the space use metrics across individuals in each group increased (Figs S2,S3).

We quantify changes in individual behavior using the correlation coefficient of behavioral metrics across experimental periods. This shows that individuals have consistency in space use and number of contests, as demonstrated by the generally positive correlations during the entire observation period (Fig 3A). At the beginning of Period 3, the correlation coefficients for all metrics except time at feeder decreased. However, in subsequent measurements in Period 3, the correlation coefficients for the space-use metrics of distance from wall, territory size, and top of nestbox then increased and remained high during Exps 11 and 12. This suggests that individuals adopted new behavioral routines that persisted after the Period 3 re-grouping.

Fig 3B shows that while there is consistency from one Exp to the next, small behavioral shifts over time can accumulate. Moreover, we see that the re-groupings facilitated changes in behavior. This is demonstrated by the sharper decrease in the correlation of behavioral metrics with Exp 7 in Periods 2 and 3 compared to that in Period 1 preceding Exp 7. In particular, while the correlation coefficient during Exps 11 and 12 during Period 3 showed a high correlation in space-use metrics (Fig 3A), the correlation of Period 3 measurements with Exp 7 remained low (Fig 3B). This indicates that the new behavioral routines of Period 3 differed from those of Period 1.

Metrics for group social structures

With the pairwise approach-avoidance interaction matrices for each Experiment, we use multiple common metrics to characterize different aspects of group social structure and

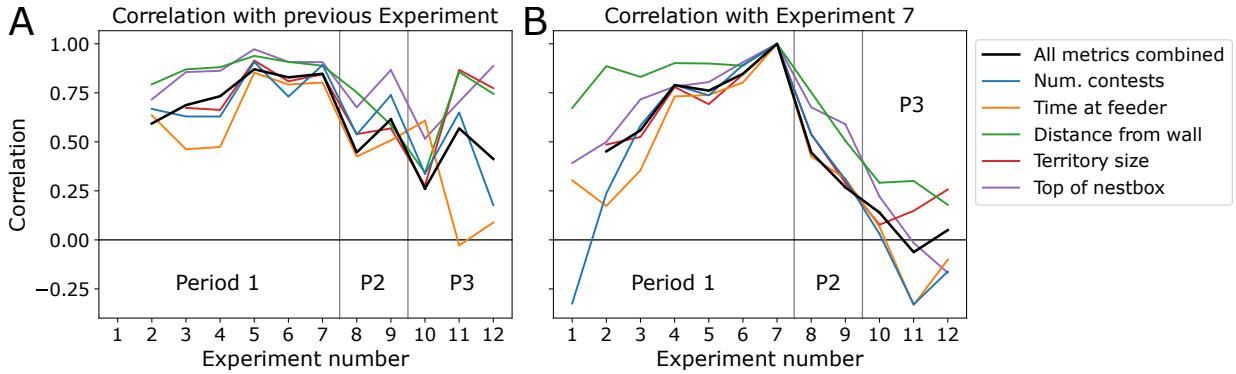

Fig 3. Correlation of space use metrics and number of contests over time. (A) Correlation with the previous experiment. (B) Correlation with Exp 7 (the last measurement in Period 1). The correlation is calculated across all rats and experimental periods with respect to a particular metric.

an individual's placement in this structure. In this section, we first describe these metrics to demonstrate what they represent, and in the following section, we use them to quantify group social structures.

We use the 5 metrics shown in Fig 4 to characterize hierarchical structures. The Elo score steepness (ESS) is a measure of the spread of the distribution of Elo scores across the group. It is calculated by converting the Elo score to a success probability, summing normalized values across group members, and calculating the slope of a linear regression fit to the resulting values [53]. The David's score steepness (DSS) (often referred to simply as hierarchy 'steepness', or 'classic steepness' [50, 53]) is calculated as the slope of a linear regression fit to the normalized David's scores among group members [55]. Individual local reaching centrality (LRC) uses the directed network of excess pairwise contest outcomes (positive entries for rats in a pair that dominated more contests, and zero for the other rat - see Methods) in order to assign higher scores to individuals in higher positions within a group hierarchy. For an unweighted directed network, LRC is the fraction of nodes reachable by any given node; a generalization of the metric accounts for weighted connections [43]. Global reaching centrality (GRC) is the average difference of nodal LRC with that of the highest LRC of any node in the graph, and a higher GRC indicates a more hierarchical network [43].

The directional consistency index (DCI) is the fraction of contests dominated by the more dominant individual of each pair, with 1 corresponding to perfect predictability in the outcome of a pairwise contest (i.e. one individual is always dominant), and 0 representing an exchange of approach-avoidance outcomes (i.e. each individual dominates the same number of contests) [50, 56]. The triangle transitivity index (TTRI) is the fraction of triad relationships that show transitivity in pairwise contest dominance outcomes (i.e. if $a \rightarrow b$ and $b \rightarrow c$, then $a \rightarrow c$ for a transitive triad) [51, 57]. Fig 4 shows idealized networks in order to illustrate what these metrics represent.

Group social structures

We find that groups differ in their social structure, but within-group structure shows consistency when group membership remains unchanged. We compare Period 1 and Period 3 group social structures because the associated experimental periods all featured groups of 7 rats. In general, we saw larger and more consistent differences between the groups in Period 3 in comparison to those in Period 1.

In Period 1, in contrast to space use, which showed clear strain-based differences (Fig S2), we do not see clear differences between strains A and B in terms of overall

	Connected hierarchy	Line	Layered hierarchy	Layered-half	Non-transitive	Single dominant	Single out	Symmetric
ESS	**1.00**	**1.00**	0.81	0.79	0.65	0.59	0.59	*0.55*
DSS	**0.99**	0.08	0.41	0.32	0.12	0.11	0.11	*0.00*
GRC	0.58	0.58	0.78	**1.00**	0.00	**1.00**	*0.03*	-
DCI	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	*0.00*
TTRI	**1.00**	-	**1.00**	**1.00**	*0.00*	-	-	-

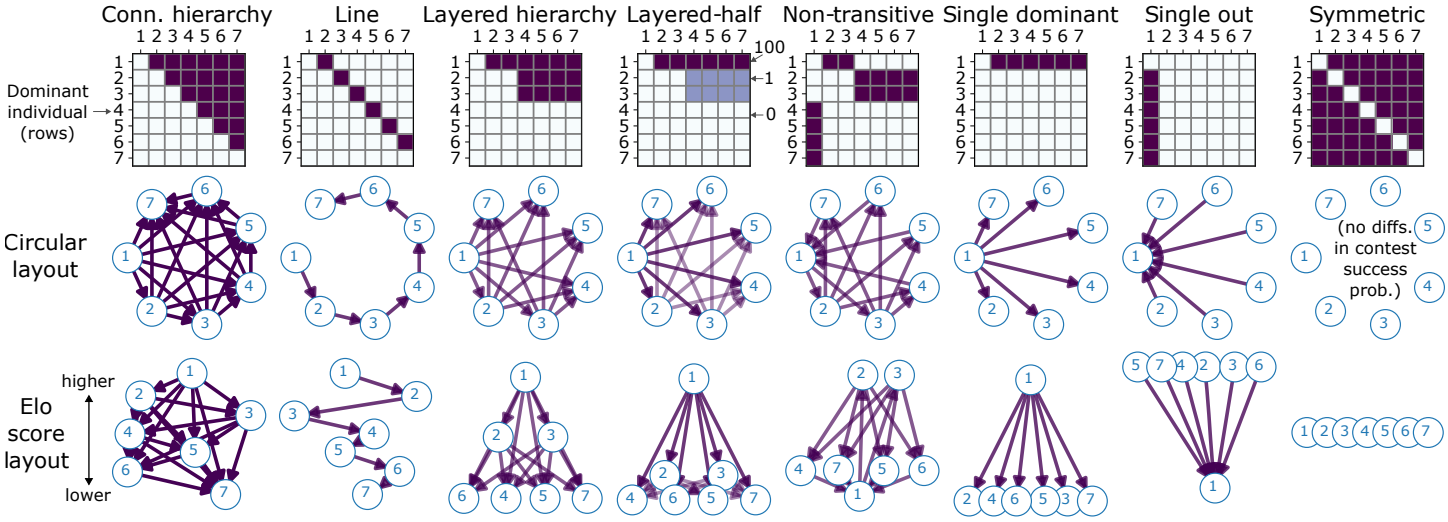

Fig 4. Idealized networks and group social metrics. The table at the top shows the scores calculated: Elo score steepness (ESS), David's score steepness (DSS), global reaching centrality (GRC), directional consistency index (DCI), and triangle transitivity index (TTRI). Note that the GRC is not defined for the symmetric network, and the TTRI is not defined for networks that do not contain any dominance triads. The different idealized networks have 7 nodes, and individual entries are either 100 or 0 (for the layered-half network, 100, 1, and 0 are used). The connected hierarchy network has a non-symmetric structure. The line network has a single "line" of pairwise interactions, where each individual only interacts with one other. The layered hierarchy network has a single individual who dominates all others and two other sub-dominant individuals who only dominate the four others below them. The layered-half network has the same structure but lower values for the sub-dominant individuals. The non-transitive network has individual 1 dominating 2-3, 2-3 dominating 4-7, but 4-7 dominating 1. The single dominant network only has contests with the dominant individual. The single-out network only has contests with the subordinate individual. The symmetric network has equal contest dominance probability among all pairs. In the table, the highest score for each metric (rows) is in bold, and the lowest score is in italics. The ESS is highest for the connected hierarchy and line networks, the DSS is highest for the connected hierarchy network while low for the line network, and the GRC is highest for the layered-half (i.e. structured but nonlinear hierarchy) and single dominant networks. In addition to matrix plots, each network is visualized by showing connections in the direction of the more to less dominant individuals in each pair (note that no connections are shown for the symmetric network because, in this case, there are no differences in contest dominance probability). We show both a circular layout and a layout based on Elo scores, where individual nodes with higher Elo scores are shown higher up on the y-axis.

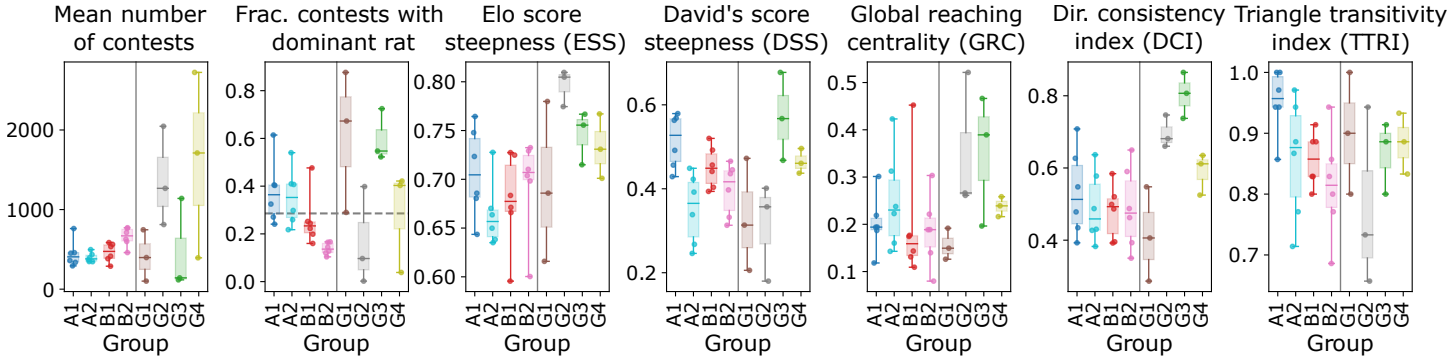

Fig 5. Measures of group social structure. The mean number of pairwise contests and the fraction of total contests with the most dominant rat are shown in addition to the hierarchy-related metrics of Elo score steepness (ESS), David's score steepness (DSS), global reaching centrality (GRC), directional consistency index (DCI), and triangle transitivity index (TTRI). The fraction of contests with the dominant rat (where "dominant" is defined as the individual with the highest Elo score) is analogous to the measure of "despotism" used in other work [31]. The metrics are calculated for each Experiment and are shown as boxplots for each Period 1 and Period 3 group. See also Fig S7 for values for each group over time.

group social structure. While the fraction of contests with the dominant rat was higher for the A groups in comparison to the B groups, other metrics do not show large or consistent differences (Fig 5). Within Period 1, each group showed a relatively strong consistency in the social structure over time, with the exception of Exp 6 for group $A2$, where a large change in the individual rank ordering in the social structure of the group took place (Fig 6A-B).

In Period 3, $G1$ and $G3$ each had a single consistent dominant individual, $G2$ had ongoing changes in social structure, and $G4$ had a stable hierarchy but with ongoing contests. All groups had consistency in structure, but the correlation of individual scores with the groups was higher for groups $G1$, $G3$, and $G4$ in comparison with group $G2$. (Fig 6C-D).

Compared to other Period 3 groups, $G1$ and $G3$ had a relatively low number of contests and a high fraction of contests with the dominant individual. Each of these groups had a single individual that was consistently ranked as most dominant (see Fig 6C). However, in comparison to $G1$, $G3$ had on average higher DSS, ESS, and GRC. This and the higher DCI index suggests that group $G3$ had a steeper hierarchical structure than group $G1$.

In contrast to groups $G1$ and $G3$, group $G2$ did not have a single individual that remained dominant during each Experiment. Group $G2$ had many contests, the lowest fraction of contests with the dominant rat, and the lowest transitivity (TTRI) of the groups. This and the lower correlation coefficient in Elo scores compared to other Period 3 groups suggest an ongoing struggle for position within the social network where ongoing contests maintained pairwise relationships. However, we note the differences obtained in the hierarchy steepness measures for group $G2$: the David's score steepness suggests weak hierarchical structures, while the Elo score steepness and GRC suggest steeper hierarchies. Overall, the inconsistency between the various metrics of hierarchy for group $G2$ suggests the absence of a distinct hierarchical structure.

Group $G4$ had similar patterns of metrics to Group $G3$, but with several distinct differences: these include lower magnitudes of ESS, DSS, GRC, and DCI, a lower fraction of contests with the dominant individual, and overall many more contests (although the mean number of contests decreased dramatically from Exp 11 to Exp 12 – see Fig S7). With this, we can describe $G4$ as having a middle-steep hierarchy that was maintained by many ongoing contests among pairs. This differs from $G1$ and $G3$, where

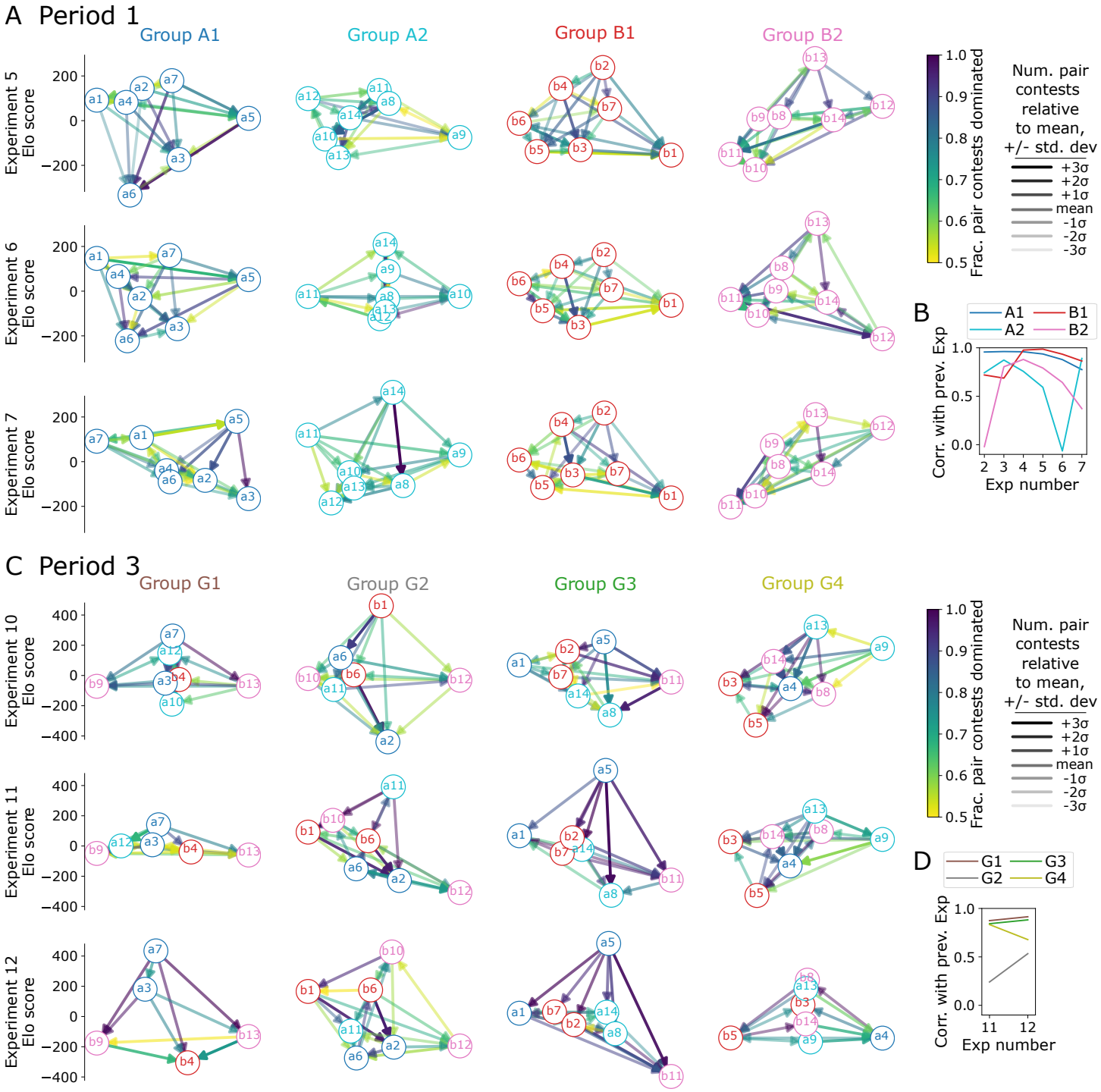

Fig 6. Social structure network visualizations. (A,C) A visualization of group networks during (A) the last 3 experiments of Period 1, and (C) Period 3. Columns correspond to different groups and rows to each experiment. The position of each individual on the y-axis is set according to their Elo score. The direction of each connection indicates which individual dominated more contests in the pair (e.g. a connection $a4 \rightarrow a6$ indicates that $a4$ more often displaced $a6$ than vice versa), the color indicates the fraction of contests dominated, and the transparency is proportional to the total number of pairwise contests relative to the mean for that group and experiment. (B,D) The correlation of individual Elo scores with the previous experiment, for (B) Period 1 groups, and (D) Period 3 groups.

the high fraction of contests with the dominant individual suggests that the hierarchy was maintained mostly by these contests.

During Period 3, the area available to each group was changed during Exps 11 and 12 by moving the compartment borders that separated the groups. In Exp 11, $G1$&$G4$ had a larger area and $G2$&$G3$ a smaller area. Exp 12 had these sizes switched. These manipulations did not have a consistent effect on space use or group social structure metrics (Fig S8).

Individual social rankings, changes over time, and body mass

We found that previous social status in Period 1 did not predict an individual's placement in the new group social structures of Period 3 (Fig 7A). For example, the most dominant rats at the end of Period 1 did not remain the most dominant during Period 3. While Fig 7A shows the correlation of Elo scores from the end of Period 1 to the beginning of Period 3, we note that the other individual metrics of fraction of contests dominated, David's score, and local reaching centrality showed similar trends (Fig S5).

Because social ranking can be used to regulate access to resources such as food, we further examine the relationship between social rank and weight gain/loss. We compare average individual social rankings to weight gain or loss during Period 3. At the beginning of Period 1, all rats were young and gaining weight. However, by the end of Period 1, the average weight gain from the previous experiment was small, and not all rats were still gaining weight. When the groups were merged in Period 2, the average change in body mass (Δ body mass) continued to decrease and was negative for the last experiment of Period 2 and the first experiment of Period 3. Specifically, in the new groups of Period 3, the variance in the distribution of Δ body mass increased, with one rat (rat $a10$, which was subsequently permanently excluded from the experiment) losing nearly 100g relative to the previous period (Fig 7B).

Fig 7C shows that although the two most dominant rats during Period 3 gained the most weight (rats $a7$ and $a5$ from groups $G1$ and $G3$, respectively), average social dominance rank was not a robust general predictor of body mass changes across all rats. Including these dominant rats, the relationship between the average Elo score during Period 3 and the change in body mass during Period 3 has a significant correlation ($r = 0.46$, $p = 0.018$). However, this result is not robust: if these two rats are removed, the correlation drops greatly to $r = 0.15$ and is no longer significant ($p = 0.47$). This demonstrates the complex relationship between dominance and body size in rodent social groups [58, 59]. While there was likely a feedback regarding social rank and body mass for the two dominant rats in groups G1 and G3, respectively, it is difficult to link weight gain/loss to social rank in a general sense.

Individual metrics compared to behavioral assays

We used individual and pairwise assays performed after the group experiments to test the behavior of each rat. The individual black and white box, canopy, and elevated plus-maze results were used to define a composite boldness score. The pairwise social tests with an unfamiliar rat were used to define social interest and self grooming scores (see Methods; Figure S10).

In general, we find low and/or inconsistent relationships between behavior in groups and behavior in the assays (Figure 8A). This is shown by the comparison of behavioral metrics from the last experiment in Period 3 with the composite individual boldness, social interest, and self grooming scores. In particular, the social interaction metrics measured in a group setting, including the number of contests and the Elo score, do not exhibit consistent or significant correlations with the pairwise social interest score.

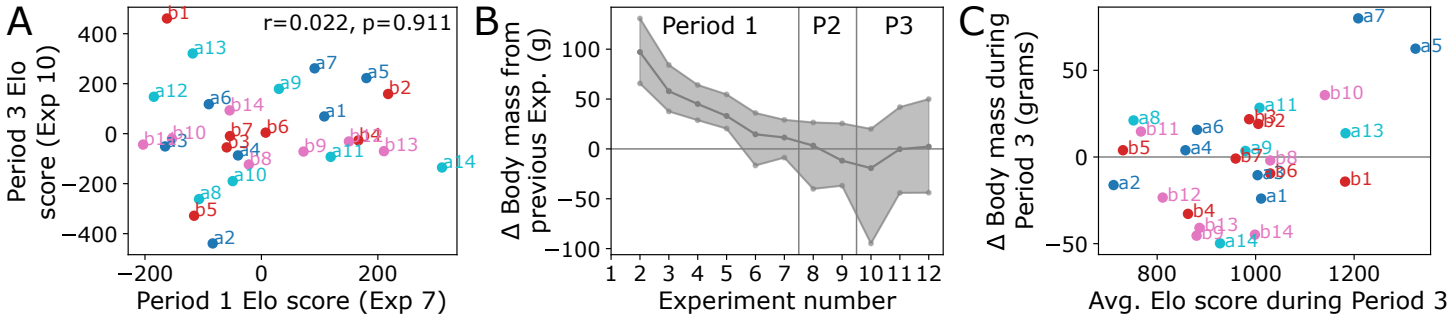

Fig 7. Period 1 to Period 3 individual rankings, body mass changes and social ranking. (A) Comparison of individual rat ranking metrics at the end of Period 1 (Exp 7) to those at the start of Period 3 (Exp 10), after the new groups were formed. (B) Distribution of body mass changes over time for all individuals. The middle line is the mean, and rugged curves indicate the maximum/minimum across all individuals. (C) Average Elo score during Period 3 compared to body mass change during Period 3 for each rat. Change in body mass during Period 3 is calculated as the body mass in Exp 12 minus body mass in Exp 10.

However, we do see positive correlations for the boldness score compared to the related metrics of distance from wall and territory size. We also find that strain and group membership do not consistently predict differences in individual test scores (Figure 8B).

The social interest and self grooming scores shown in Figure 8A-B were obtained via pairwise behavioral assays performed with an unfamiliar rat. We repeated these tests with a second unfamiliar rat in order to test repeatability (tests were also performed with a familiar rat from the respective Period 3 groups - see Figure S11). The composite scores from the tests with the second unfamiliar individual show a low correlation with the scores obtained with the first unfamiliar individual (Figure 8C). These comparisons support the general conclusion that behavioral response in pairwise assays has limited predictive power in relation to social behavior in a more complex environment and group setting.

Discussion

We utilized automated tracking techniques to describe how rat groups develop and maintain a dynamic social structure over time, as well as how the social structure changes after regrouping. Across successive Experiments, we observed a general consistency in the behavior of both individuals and groups. However, considering longer periods of time across multiple Experiments, we observed that the gradual accumulation of small changes can result in substantial behavioral changes over these longer time scales. In addition, when the group composition was altered, we observed accelerated changes in behavior. Multiple metrics, including the Elo score, David's score, and reaching centrality, were employed to describe the overall hierarchical structure of each group as well as individual social rankings; these metrics revealed both similar and dissimilar aspects of the structure. Different groups can vary a lot in their structure, and in particular we found different and distinct social structures in the newly formed groups of Period 3. We found that an individual's new position in the social hierarchy cannot be predicted based on their prior status when group composition changes. While conventional individual assays (including the elevated cross, canopy, and other tests) produce consistent test results, we found that these measures have little correlation with individual behavior in a group setting. Moreover, we found low repeatability in the scores measured with standard social test assays by performing the same test with different individuals; this also contributes to why behavior assays have little correlation

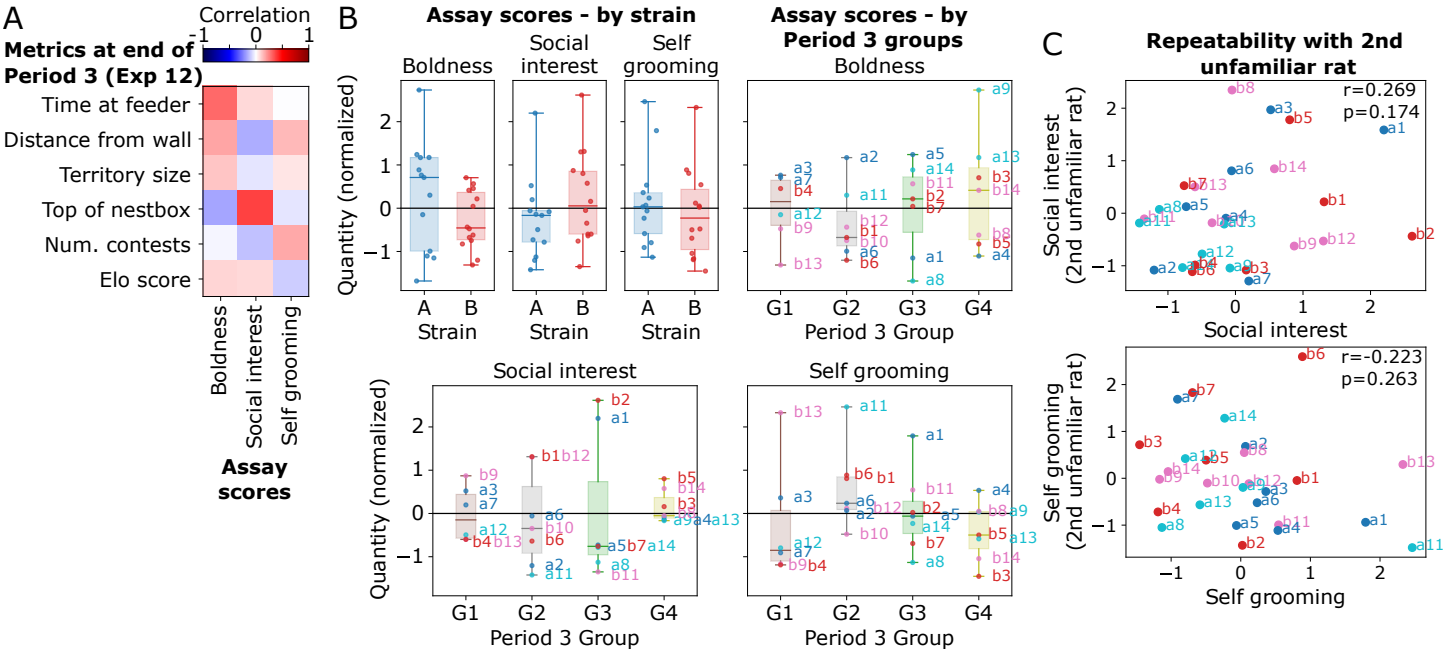

Fig 8. Behavioral metrics at the end of Period 3 compared to assays. (A) Correlation of space use and social behavioral metrics from the final Experiment in Period 3 (Exp 12) with individual assay scores. The colors indicate correlation values. (B) Individual score distributions according to strain (left), and by Period 3 group membership (right). Scores are normalized by the mean and standard deviation of values measured for all rats. (C) Comparison of social interest and self grooming scores calculated from tests with a first unfamiliar rat (x-axes; values shown in A-B), with scores calculated from tests with a second unfamiliar rat (y-axes). See also Figure S11

with group behavior.

At the beginning of Period 1, the rats were still juveniles. Social interactions, particularly those related to aggression and dominance, are known to develop over time [11, 22, 60]. Our observations are consistent with this, in particular because we observed an increase in the number of wounds from fighting at the end of Period 1 (Figure 2). It is likely that in addition to group composition and interactions, development stage also influenced the number of contests and the differences in social structure from Period 1 to Period 3. In a natural population, groups consist of individuals of different ages [22]. Targeted group mixing experiments – for example, with both juveniles and adults in the same group – could be used to ask how these effects interplay to generate overall emergent group social structures.

Behavioral assays are often used to quantify the behavior of rodents, and many new tools are being developed for both individual and social tracking and behavioral scoring [61–66]. Tests that have been used to quantify social behavior in rodents include, for example, reciprocal social interaction, social approach partition, social preference, social transmission of food preference, food allocation, and reciprocal cooperation [13, 23, 67]. Most of these tests rely on creating artificial situations in order to measure the corresponding behavioral outcome. It is an open question as to whether the social behavior measured in such tests can predict the social behavior under more natural conditions that include complex social interactions [68]. Our comparison with the results of pairwise social interest behavioral assays highlights the need for further work in this area. We also note that while our group-living experiments provided space for complex environmental and social interactions, the conditions in the experiments were still much different from those that rats experience in the wild. In this respect, our methods are similar to recent studies with mice, such as experiments with groups in the "social box" [29–31]. These experimental setups can, therefore, be described as semi-natural. A unique aspect of our study is the extended observation period, which allowed us to examine not only group social structure but also its change over time.

Long-term studies to examine group behavior may be an essential component to include in translational research applications, for example, in the testing of psychotherapeutic drugs to treat social anxiety [24, 25]. In particular, an important topic for future work is to establish standardized and reproducible tests and measures that are properly representative of a full range of social behavior [27, 69]. Another important application is understanding the neural basis of social interactions. The great majority of our neuroscientific understanding of social behavior comes from dyadic interactions and reduced forms of social interactions [26, 70–72]. In light of the expanding interest in the neuroscience of natural social behavior [68], going beyond basic social testing paradigms and static characterizations of social hierarchy lends the opportunity to unravel a richer repertoire of neural mechanisms.

We note that while our social behavior analysis was used with video tracking data, it could be applied to other types of data, for example, data derived from markerless tracking methods, motion capture or QR-code tracking. Future work can continue to expand on methods in this area, for example, including more detailed posture data, which can be used to describe social interactions in more detail [73–76]. Systems in this direction have already been developed for use with mice [63]. While we used approach-avoidance interactions to define pairwise contests, we note that a more detailed approach could use a combination of multiple behavioral interactions in order to define contest "winners" and "losers" (e.g. [77]). Another area for future work is testing the functional consequences of group composition and social structure on individual or group performance, for example, with respect to search [3].

Author contributions

T.V. and M.N. conceived the idea and designed the project. M.N., G.V., D.Á., and E.K. conducted the experiments and collected the data. E.K. performed the analysis of the individual assays. G.V. designed and wrote the software framework for tracking the individuals. M.N., G.V., and V.T. designed and performed the analysis of the behavior within the groups. J.D.D., M.N., and A.E.H designed the analysis of the group-level data. J.D.D. performed the analysis of the group-level data. J.D.D., M.N., and A.E.H wrote the initial draft of the manuscript, and all authors revised the manuscript.

Acknowledgments

We thank all the people who helped during the experimental work, especially Valéria Németh, Gergő Somorjai, Zsuzsa Ákos, Katalin Schlett, Péter Urtz, and András Péter (for Solomon coder). We acknowledge funding from Eötvös Loránd University and the Max Planck Institute of Animal Behavior. This research was partially supported by grants to T.V. establishing the MTA-ELTE Statistical and Biological Research Group, including NKFI Office Grant OTKA SNN 139598. Other sources of funding include the Hungarian Academy of Sciences (a grant to the MTA-ELTE 'Lendület' Collective Behaviour Research Group, grant number 95152 and the Companion Animal Research Group, grant no. PH1404/21), and the DFG Centre of Excellence 2117 "Centre for the Advanced Study of Collective Behaviour" (ID: 422037984). M.N. was partially supported by the Hungarian National Research, Development and Innovation Office (grant no. K 128780) and by Isaac Newton Institute for Mathematical Sciences, University of Cambridge. J.D. was partially supported by the Heidelberg Academy of Sciences Young Scientist (WIN-Kolleg) program.

Methods

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Mate Nagy (nagymate@hal.elte.hu).

Materials availability

The study did not generate new unique reagents.

Data and code availability

Trajectory and behavioral metrics data are available through Zenodo at doi.org/10.5281/zenodo.7615468. Recorded video sequences were analyzed offline with a custom-written software to obtain individual positions and orientations (source code available at github.com/vasarhelyi/ratognize), as well as trajectories and metrics (source code available at github.com/vasarhelyi/trajognize). For details, see SI Methods of [3]. The scripts to run analyses included in this paper are available at github.com/jacobdavidson/ratsocialgroups.

Experimental model and subject details

Subjects

We used 28 Wistar male rats from 2 inbred strains (14 Crl:WI BR and 14 HsdBrlHan:WIST, 7 litters/strain, 2 individuals/litter; ordered from Toxi-Coop Zrt, Hungary) in this study. The rats arrived on 24 May 2011 at an age of 6 weeks. Rats were separated into 4 same-strain groups (i.e. the Period 1 groups), each containing 7 rats from different litters, and were housed and tested together.

Each rat was marked with a unique 3-color barcode on its back using nontoxic "Special Effects" hair dye in 5 distinctive colors (Red: Nuclear Red, Orange: Napalm, Green: Sonic Green, Blue: Londa color 0/88, Purple: 4 units Atomic Pink and 1 unit Wild Flower). These codes were applied/renewed every 3 weeks.

Ethical guidelines

The procedures comply with national and EU legislation and institutional guidelines. The experiments were performed in the animal facilities of the Eotvos University, Hungary, and in accordance with Hungarian legislation and the corresponding definition by law (1998. évi XXVIII. Törvény 3. §/9. — the Animal Protection Act), which states that noninvasive studies on animals bred for research are allowed to be performed without the requirement of any special permission (PE/EA/1360-5/2018).

Method details

Experimental conditions and monitoring

Animals were housed in 4 compartments (sized 100 x 125 x 100 cm^3) with polypropylene covered wooden walls and sawdust changed weekly on a tiled floor. Their room was kept at a controlled temperature of 21 $\pm$ 2°C, and with controlled light conditions featuring a daily cycle with 13h/11h dark/light. The dark (active) period was from 6am to 7pm with illumination at floor level ∼3-4 lux; the light period followed

from 7pm to 6am of the following day with illumination of 300 lux. We video recorded the compartment 24/7 using a low-light sensitive camera fixed to the ceiling (Sony HDR-AX2000, 2.9 $\times$ 1.8 m^2 field of view, 1920 $\times$ 1080 resolution, 25 fps de-interlaced). Rats had *ad libitum* access to water and a shelter (nestbox), and access to food based on a fixed schedule for automated feeding. The feeding schedule followed a weekly cycle: 3 days (Sat, Sun, Mon) access to food 3 times for 1 hour (at 6am, noon, and 6pm); 3 days (Tue, Wed, Thu) access to food 2 times for 1 hour (at 6am and 6pm); and 1 day (Fri) access to food *ad libitum* between 6am and 7pm). The housing compartments were cleaned once a week.

We measured the weight of each individual three times a week (Mon, Thu, Fri; between 5pm and 6pm) and counted fresh wounds. Classification of wounds followed a predefined but arbitrary scale ranging from 0 to 6, which used the following conventions: 0 $\rightarrow$ no wounds received in the last 3 days, 1 $\rightarrow$ 1 fresh wound, 2 $\rightarrow$ 2 fresh wounds, 3 $\rightarrow$ few (3-4) fresh wounds, 4 $\rightarrow$ more than few (5-7) wounds, 5 $\rightarrow$ many (8-12) wounds, and 6 $\rightarrow$ even more wounds (13 or more). When a rat was seriously injured, we temporarily removed it until it was recovered. This happened on one occasion: $a12$ was taken out from 5 to 27 Jan. 2012. One rat ($a10$) was permanently removed due to injuries and weight loss at week 31 (19 Dec 2011) at the age of 37 weeks.

Individual and social interaction tests

These tests were performed at the end of the group measurement period, and included a total of 27 male rats at the age of 44 weeks. Subjects participated a test battery consisting of seven subtests examining fear-related and social behaviors in the following order (see descriptions below): black and white box, canopy, elevated plus-maze, social interaction test with outgroup conspecific, and social interaction test with in-group conspecific. The illumination during these tests was set according to the dark period mentioned above. Body weights were 480$\pm$70g (mean$\pm$SD) at the time of these tests. The behavioral tests were conducted on three consecutive days between 10 a.m. and 6 p.m. (26 to 28 Feb 2012). The apparatuses were constructed from plastic sheets and cleaned between tests. Depending on the test, either automated analysis was used to obtain trajectories or the behavior was coded by observers using the Solomon Coder software (beta 19.08.02). To ensure inter-observer reliability, pairs of observers overlapped in 20% of the behavioral tests they scored. Using this overlap, the inter-observer reliability was calculated using the intraclass correlation coefficient (ICC) for all variables except the video frame variables, and we found all observations to be reliable (ICC $>$ 0.9).

1. Black and white box. As described in Ramos et al. [78], the apparatus had a black and a white compartment (each sized 27x27x27). The white compartment was strongly illuminated by a white bulb ($\sim$825 lux), while the black compartment was illuminated with a red bulb ($\sim$90 lux; Figure S9A). The bulbs were positioned 37 cm above the apparatus floor. Each rat was initially placed in the center of the white compartment in a direction facing the opposing black compartment, and behavior was then recorded for 5 minutes. We tallied the number of video frames when the animal was in the 1) white compartment, 2) black compartment, or 3) at the border of the two areas and thus could not be clearly assigned (labeled as "both" in the data).

2. Canopy. The apparatus consisted of a circular platform (104 cm in diameter) and a canopy (semitransparent red Plexiglas of 70 cm diameter) 10 cm above the platform (Figure S9B). The mean illumination was 90 lux under the canopy and 400 lux outside of the canopy. At the beginning of the test, the animal was placed under the canopy. The test lasted for 5 minutes. We counted the number of video frames when the animal was 1) under the canopy, 2) in the exposed zone.

3. Elevated plus-maze. Based on Ramos et al. [78], the apparatus had four

elevated arms (66 cm from the floor), 45 cm long and 10 cm wide (Figure S9C). Two closed arms enclosed by a 50 cm high wall were located on opposing sides, and two open arms on the other two sides; the wall structure led to different illumination, with 25 lux for the closed arms and 65 lux for the open arms. The central platform (10× 10 cm) connected the four arms to allow access to any. Each rat was first placed in the central platform facing an open arm, and subsequently, behavior was recorded for 5 minutes. To describe behavior, we counted the number of video frames in which the animal was in the 1) closed arms, 2) open arms, and 3) central platform (labeled as "both" in the data).

4. Social interaction test with an unfamiliar (out-group) conspecific. In an open field arena, we placed an unfamiliar adult male next to a focal rat that had been part of the long-term experiment. Two different unfamiliar rats were used for each Period 3 group (i.e. $G1$ rats were tested with unfamiliar rats 1 and 2, $G2$ rats with unfamiliar rats 3 and 4, etc.). The test apparatus was made out of glass, with a green floor of 74 × 74 cm and transparent walls (∼40 cm high; Figure S9D). The unfamiliar rats were significantly younger and smaller than the focals (12-weeks-old and 360±20g (mean±SD)). We recorded the behavior for 10 minutes. The test was repeated with other rats (unfamiliar and familiar) after a break that lasted for on average 75±42 min. but a minimum of 35 min. The coded behavior included the following: duration of bipedal orienting stance (%), duration of self grooming (%), duration of exploration, duration of sniffing non-genital body parts (%), duration of sniffing the genitals of the partner (%), number of steps on the partner, number of fights. The coded parameters were stored separately from the trials with unfamiliar rats 1 and 2 for each focal rat.

5. Social interaction test with a familiar (in-group) conspecific. The social interaction was also performed with a familiar conspecific, chosen as a random groupmate from their Period 3 group. Each individual was measured with four randomly selected members from their group. The coded variables were the same as above, but they were averaged over the trials for each focal rat.

Quantification and statistical analysis

Data processing and behavioral metrics

We calculated metrics of space use from the individual trajectory data for each rat. Time at feeder is the fraction of total time spent at the feeder during nightlight (active period). Distance from wall is the average distance from the walls during nightlight. Territory size is the area of an individual's space-use heatmap during all times (number of bins where it was detected more than 10 frames per day, calculated using bins with a linear size of ∼ 2 mm) (see Figure 1C), normalized by total number of bins and frames counted. Top of nestbox is the fraction of total time spent on top of the nestbox/shelter area during nightlight.

An approach-avoidance (AA) contest event was defined for a given pair of individuals ($i \neq j$) if, for a 0.4s long time window, the time-averaged dot product of i's velocity ($v_i(t)$) and the normalized relative direction vector pointing from i to j – a unit vector $\hat{d}_{ij}(t) = d_{ij}(t)/|d_{ij}(t)|$, where $d_{ij}(t) = x_j(t) - x_i(t)$ is the relative position – were within predetermined thresholds for both individuals. The thresholds used were $AA_{ij} = \langle (v_i(t) \cdot \hat{d}_{ji}(t) \rangle_t > 0.8$ for the approacher, and $AA_{ji} < -0.5$ for the avoider. In addition, we used the requirement that i and j were within 40 cm of each other ($|d_{ij}(t)| \leq d_{max} = 40$ cm) and both i and j were moving at speeds of at least 0.25 m/sec ($|v(t)| \geq v_{min} = 0.25$ m/sec).

We use the approach-avoidance contest network to quantify the social interaction structure in each group. The values A_{ij} are the number of times rat i dominated approach-avoidance contests with rat j. Using this, the number of contests rat i

dominated is $w_i = \sum_j A_{ij}$, and the number of contests lost is $l_i = \sum_j A_{ji}$. The fraction of contests dominated for rat i is then

$$f_i = \frac{w_i}{w_i + l_i}. \tag{1}$$

Reaching centrality is calculated using the normalized network of excess wins, W_{ij}. This network has positive entries for the rat in a pair that dominated in more contests and zero for the other rat, and is determined as

$$W_{ij} = \frac{1}{Z} max\left(A_{ij} - A_{ji}, 0\right), \tag{2}$$

where Z is a normalization factor, which we define so that the maximum entry of of W_{ij} is equal to 1. This network is then provided as input to the `networkx` function `local_reaching_centrality` to calculate the local reaching centrality (LRC) for each individual, and to the function `global_reaching_centrality` to calculate global reaching centrality (GRC) for the group. Note that we set the flag `normalized=False` for calling these functions, because we use the definition in Eq 2 where W_{ij} is already normalized. With this, the LRC and GRC scores are in the range of 0 to 1.

We used the `EloRating` package [79] to calculate the individual David's score, David's score steepness (DSS), directional consistency index (DCI), and triangle transitivity index (TTRI). For the individual Elo scores, we used the `EloChoice` package [79], which has an improved and more efficient implementation of the randomization of interaction sequences used to calculate the Elo score. We used the `EloSteepness` package [80] to calculate the Elo score steepness (ESS). These R packages were integrated into our Python-based analysis code using `rpy2`.

Individual and social interaction assays

We used principal component analysis (PCA) to define the boldness, social interest, and self grooming scores for each rat from the individual and pairwise assays.

Boldness score. The 8 videoframe variables calculated from the individual tests (black-and-white box, elevated plus-maze, and canopy test) were used to define the boldness score, which reflects the time spent in exposed portions of an unfamiliar environment. We converted each frame count to a fraction of test time and normalized the input variables before applying PCA. The first component explains 44.7% of the variance, and positive projections onto this component represent more time in open areas (Figure S10A). We used the projection of each rat onto the first component as the "boldness" score. For comparison, we also calculated fractions of open-area time for each test: fraction in the white area during black and white black box test (BWB-whitefrac = BWB-White/(BWB-White + BWB-Black)), fraction of time in open during the elevated cross test (ElevX-openfrac = ElevX-Open/(ElevX-Open + ElevX-Closed)), and fraction of time out during the canopy test (Canopy-outfrac = Canopy-OUT/(Canopy-Out + Canopy-Under)).

Social interest and self grooming scores. We applied PCA to measures from the pairwise social interaction tests and used results to define composite scores related to social interest and self grooming. The variables included are duration of sniffing genitals (%), duration of sniffing nongenital body parts (%), duration of bipedal orienting stance (standing up) (%), number of steps on the partner, number of mating attempts, number of fights, duration of exploration (%), and duration of self grooming (%). The components shown in Figure S10B were determined using data from the first test with an unfamiliar rat; the scores for the other tests with a familiar rat and a second unfamiliar rat were calculated by projecting the associated variables onto these components. The first PCA component represents interaction with the other rat, with

positive projections indicating more interactions. We use this component as the "social interest" score. The second PCA component is weighted most strongly by self grooming (positive) and exploration (negative) – we use this component as the "self grooming" score.

Figure S11 shows a comparison of the scores, including the boldness score and measures from the individual tests, and the social interest and self grooming scores from the first test with an unfamiliar rat as well as tests with a familiar rat and a second unfamiliar rat.

References

1. Vishwakarma M, Russo JD, Probst D, Schwarz US, Das T, Spatz JP. Mechanical interactions among followers determine the emergence of leaders in migrating epithelial cell collectives. Nature Communications. 2018;9(1):3469. doi:10.1038/s41467-018-05927-6.

2. Davidson JD, Arauco-Aliaga RP, Crow S, Gordon DM, Goldman MS. Effect of interactions between harvester ants on forager decisions. Frontiers in Ecology and Evolution. 2016;4:115. doi:10.3389/fevo.2016.00115.

3. Nagy M, Horicsányi A, Kubinyi E, Couzin ID, Vásárhelyi G, Flack A, et al. Synergistic Benefits of Group Search in Rats. Current Biology. 2020;30(23):4733–4738.e4. doi:10.1016/j.cub.2020.08.079.

4. Bond RM, Fariss CJ, Jones JJ, Kramer ADI, Marlow C, Settle JE, et al. A 61-million-person experiment in social influence and political mobilization. Nature. 2012;489(7415):295–298. doi:10.1038/nature11421.

5. Kaburu SSK, Newton-Fisher NE. Egalitarian despots: hierarchy steepness, reciprocity and the grooming-trade model in wild chimpanzees, Pan troglodytes. Animal Behaviour. 2015;99:61–71. doi:10.1016/j.anbehav.2014.10.018.

6. Nagy M, Vásárhelyi G, Pettit B, Roberts-Mariani I, Vicsek T, Biro D. Context-dependent hierarchies in pigeons. Proceedings of the National Academy of Sciences. 2013;110(32):13049–13054. doi:10.1073/pnas.1305552110.

7. Maruska KP. Social regulation of reproduction in male cichlid fishes. General and Comparative Endocrinology. 2014;207:2–12. doi:10.1016/j.ygcen.2014.04.038.

8. King AJ, Douglas CM, Huchard E, Isaac NJ, Cowlishaw G. Dominance and affiliation mediate despotism in a social primate. Current Biology. 2008;18(23):1833–1838.

9. Stewart AJ, Mosleh M, Diakonova M, Arechar AA, Rand DG, Plotkin JB. Information gerrymandering and undemocratic decisions. Nature. 2019;573(7772):117–121. doi:10.1038/s41586-019-1507-6.

10. Majolo B, Lehmann J, de Bortoli Vizioli A, Schino G. Fitness-related benefits of dominance in primates. American Journal of Physical Anthropology. 2012;147(4):652–660. doi:10.1002/ajpa.22031.

11. Barnett SA. An Analysis of Social Behaviour in Wild Rats. Proceedings of the Zoological Society of London. 1958;130(1):107–152. doi:10.1111/j.1096-3642.1958.tb00565.x.

12. Grant EC. An Analysis of the Social Behaviour of the Male Laboratory Rat. Behaviour. 1963;21(3-4):260–281. doi:10.1163/156853963X00194.

13. Schweinfurth MK. The social life of Norway rats (Rattus norvegicus). eLife. 2020;9:e54020. doi:10.7554/eLife.54020.

14. Dennis EJ, El Hady A, Michaiel A, Clemens A, Tervo DRG, Voigts J, et al. Systems Neuroscience of Natural Behaviors in Rodents. Journal of Neuroscience. 2021;41(5):911–919.

15. Forkosh O. Animal behavior and animal personality from a non-human perspective: Getting help from the machine. Patterns. 2021;2(3). doi:10.1016/j.patter.2020.100194.

16. Hinz RC, Polavieja GGd. Ontogeny of collective behavior reveals a simple attraction rule. Proceedings of the National Academy of Sciences. 2017; p. 201616926. doi:10.1073/pnas.1616926114.

17. Smith ML, Davidson JD, Wild B, Dormagen DM, Landgraf T, Couzin ID. Behavioral variation across the days and lives of honey bees. iScience. 2022;25(9). doi:10.1016/j.isci.2022.104842.

18. Evans JC, Lindholm AK, König B. Long-term overlap of social and genetic structure in free-ranging house mice reveals dynamic seasonal and group size effects. Current Zoology. 2021;67(1):59–69. doi:10.1093/cz/zoaa030.

19. Pearce JM. Animal Learning and Cognition: An Introduction. 3rd ed. London: Psychology Press; 2014.

20. Phifer-Rixey M, Nachman MW. Insights into mammalian biology from the wild house mouse Mus musculus. eLife. 2015;4:e05959. doi:10.7554/eLife.05959.

21. Bale TL, Abel T, Akil H, Carlezon Jr WA, Moghaddam B, Nestler EJ, et al. The critical importance of basic animal research for neuropsychiatric disorders. Neuropsychopharmacology. 2019;44(8):1349–1353. doi:10.1038/s41386-019-0405-9.

22. Calhoun JB. The Ecology and Sociology of the Norway Rat. U.S. Department of Health, Education, and Welfare, Public Health Service; 1963.

23. Silverman JL, Yang M, Lord C, Crawley JN. Behavioural phenotyping assays for mouse models of autism. Nature Reviews Neuroscience. 2010;11(7):490–502. doi:10.1038/nrn2851.

24. Homberg JR. Measuring behaviour in rodents: towards translational neuropsychiatric research. Behavioural brain research. 2013;236:295–306.

25. Peters SM, Pothuizen HHJ, Spruijt BM. Ethological concepts enhance the translational value of animal models. European Journal of Pharmacology. 2015;759:42–50. doi:10.1016/j.ejphar.2015.03.043.

26. Kondrakiewicz K, Kostecki M, Szadzińska W, Knapska E. Ecological validity of social interaction tests in rats and mice. Genes, Brain and Behavior. 2019;18(1):e12525. doi:10.1111/gbb.12525.

27. Shemesh Y, Chen A. A paradigm shift in translational psychiatry through rodent neuroethology. Molecular Psychiatry. 2023; p. 1–11. doi:10.1038/s41380-022-01913-z.

28. Jabarin R, Netser S, Wagner S. Beyond the three-chamber test: toward a multimodal and objective assessment of social behavior in rodents. Molecular Autism. 2022;13(1):41. doi:10.1186/s13229-022-00521-6.

29. Shemesh Y, Sztainberg Y, Forkosh O, Shlapobersky T, Chen A, Schneidman E. High-order social interactions in groups of mice. eLife. 2013;2:e00759. doi:10.7554/eLife.00759.

30. Forkosh O, Karamihalev S, Roeh S, Alon U, Anpilov S, Touma C, et al. Identity domains capture individual differences from across the behavioral repertoire. Nature Neuroscience. 2019;22(12):2023–2028. doi:10.1038/s41593-019-0516-y.

31. Karamihalev S, Brivio E, Flachskamm C, Stoffel R, Schmidt MV, Chen A. Social dominance mediates behavioral adaptation to chronic stress in a sex-specific manner. eLife. 2020;9:e58723. doi:10.7554/eLife.58723.

32. Lopez JP, Lücken MD, Brivio E, Karamihalev S, Kos A, De Donno C, et al. Ketamine exerts its sustained antidepressant effects via cell-type-specific regulation of Kcnq2. Neuron. 2022;110(14):2283–2298.e9. doi:10.1016/j.neuron.2022.05.001.

33. Williamson CM, Lee W, Curley JP. Temporal dynamics of social hierarchy formation and maintenance in male mice. Animal Behaviour. 2016;115:259–272. doi:10.1016/j.anbehav.2016.03.004.

34. Lee W, Dowd HN, Nikain C, Dwortz MF, Yang ED, Curley JP. Effect of relative social rank within a social hierarchy on neural activation in response to familiar or unfamiliar social signals. Scientific Reports. 2021;11(1):2864. doi:10.1038/s41598-021-82255-8.

35. Lee W, Milewski TM, Dwortz MF, Young RL, Gaudet AD, Fonken LK, et al. Distinct immune and transcriptomic profiles in dominant versus subordinate males in mouse social hierarchies. Brain, Behavior, and Immunity. 2022;103:130–144. doi:10.1016/j.bbi.2022.04.015.

36. Johnson JA. Supplanting by olive baboons: dominance rank difference and resource value. Behavioral Ecology and Sociobiology. 1989;24(5):277–283. doi:10.1007/BF00290903.

37. Evans JC, Devost I, Jones TB, Morand-Ferron J. Inferring dominance interactions from automatically recorded temporal data. Ethology. 2018;124(3):188–195. doi:10.1111/eth.12720.

38. Gullstrand J, Claidière N, Fagot J. Computerized assessment of dominance hierarchy in baboons (Papio papio). Behavior Research Methods. 2021;53(5):1923–1934. doi:10.3758/s13428-021-01539-z.

39. Elo AE. The rating of chessplayers, past and present. New York: Arco Pub; 1978.

40. Albers PCH, De Vries H. Elo-rating as a tool in the sequential estimation of dominance strengths. Animal Behaviour. 2001;61(2):489–495. doi:10.1006/anbe.2000.1571.

41. Neumann C, Duboscq J, Dubuc C, Ginting A, Irwan AM, Agil M, et al. Assessing dominance hierarchies: validation and advantages of progressive evaluation with Elo-rating. Animal Behaviour. 2011;82(4):911–921. doi:10.1016/j.anbehav.2011.07.016.

42. Strauss ED, Holekamp KE. Inferring longitudinal hierarchies: Framework and methods for studying the dynamics of dominance. Journal of Animal Ecology. 2019;88(4):521–536. doi:10.1111/1365-2656.12951.

43. Mones E, Vicsek L, Vicsek T. Hierarchy Measure for Complex Networks. PLOS ONE. 2012;7(3):e33799. doi:10.1371/journal.pone.0033799.

44. Ozogány K, Vicsek T. Modeling the Emergence of Modular Leadership Hierarchy During the Collective Motion of Herds Made of Harems. Journal of Statistical Physics. 2015;158(3):628–646. doi:10.1007/s10955-014-1131-7.

45. Shimoji H, Abe MS, Tsuji K, Masuda N. Global network structure of dominance hierarchy of ant workers. Journal of The Royal Society Interface. 2014;11(99):20140599. doi:10.1098/rsif.2014.0599.

46. Kora Y, Salhi S, Davidsen J, Simon C. Global excitability and network structure in the human brain. Physical Review E. 2023;107(5):054308. doi:10.1103/PhysRevE.107.054308.

47. Hu F, Zhao S, Bing T, Chang Y. Hierarchy in industrial structure: The cases of China and the USA. Physica A: Statistical Mechanics and its Applications. 2017;469:871–882. doi:10.1016/j.physa.2016.11.083.

48. Beardsley K, Liu H, Mucha PJ, Siegel DA, Tellez JF. Hierarchy and the Provision of Order in International Politics. The Journal of Politics. 2020;82(2):731–746. doi:10.1086/707096.

49. Mones E, Pollner P, Vicsek T. Universal hierarchical behavior of citation networks. Journal of Statistical Mechanics: Theory and Experiment. 2014;2014(5):P05023. doi:10.1088/1742-5468/2014/05/P05023.

50. Strauss ED, DeCasien AR, Galindo G, Hobson EA, Shizuka D, Curley JP. DomArchive: a century of published dominance data. Philosophical Transactions of the Royal Society B: Biological Sciences. 2022;377(1845):20200436. doi:10.1098/rstb.2020.0436.

51. Shizuka D, McDonald DB. A social network perspective on measurements of dominance hierarchies. Animal Behaviour. 2012;83(4):925–934. doi:10.1016/j.anbehav.2012.01.011.

52. de Vries H. An improved test of linearity in dominance hierarchies containing unknown or tied relationships. Animal Behaviour. 1995;50(5):1375–1389. doi:10.1016/0003-3472(95)80053-0.

53. Neumann C, Fischer J. Extending Bayesian Elo-rating to quantify the steepness of dominance hierarchies. Methods in Ecology and Evolution. 2023;14(2):669–682. doi:10.1111/2041-210X.14021.

54. Spruijt BM, Hol T, Rousseau J. Approach, avoidance, and contact behavior of individually recognized animals automatically quantified with an imaging technique. Physiology & Behavior. 1992;51(4):747–752. doi:10.1016/0031-9384(92)90111-E.

55. de Vries H, Stevens JMG, Vervaecke H. Measuring and testing the steepness of dominance hierarchies. Animal Behaviour. 2006;71(3):585–592. doi:10.1016/j.anbehav.2005.05.015.

56. van Hooff JARAM, Wensing JAB. Dominance and its behavioral measures in a captive wolf pack. In: Man and wolf: Advances, issues, and problems in captive wolf research. Perspectives in vertebrate science, Vol. 4.. Dordrecht, Netherlands: Dr W Junk Publishers; 1987. p. 219–252.

57. Shizuka D, McDonald DB. The network motif architecture of dominance hierarchies. Journal of The Royal Society Interface. 2015;12(105):20150080. doi:10.1098/rsif.2015.0080.

58. Boreman J, Price E. Social dominance in wild and domestic Norway rats (Rattus norvegicus). Animal Behaviour. 1972;20(3):534–542. doi:10.1016/S0003-3472(72)80018-6.

59. Macdonald DW, Berdoy M, Smith P. Stability of Social Status in Wild Rats: Age and the Role of Settled Dominance. Behaviour. 1995;132(3-4):193–212. doi:10.1163/156853995X00694.

60. Bolles RC, Woods PJ. The ontogeny of behaviour in the albino rat. Animal Behaviour. 1964;12(4):427–441. doi:10.1016/0003-3472(64)90062-4.

61. Spink AJ, Tegelenbosch RAJ, Buma MOS, Noldus LPJJ. The EthoVision video tracking system—A tool for behavioral phenotyping of transgenic mice. Physiology & Behavior. 2001;73(5):731–744. doi:10.1016/S0031-9384(01)00530-3.

62. Hong W, Kennedy A, Burgos-Artizzu XP, Zelikowsky M, Navonne SG, Perona P, et al. Automated measurement of mouse social behaviors using depth sensing, video tracking, and machine learning. Proceedings of the National Academy of Sciences. 2015;112(38):E5351–E5360. doi:10.1073/pnas.1515982112.

63. de Chaumont F, Ey E, Torquet N, Lagache T, Dallongeville S, Imbert A, et al. Real-time analysis of the behaviour of groups of mice via a depth-sensing camera and machine learning. Nature Biomedical Engineering. 2019;3(11):930–942. doi:10.1038/s41551-019-0396-1.

64. Nilsson SR, Goodwin NL, Choong JJ, Hwang S, Wright HR, Norville ZC, et al.. Simple Behavioral Analysis (SimBA) – an open source toolkit for computer classification of complex social behaviors in experimental animals; 2020. Available from: `https://www.biorxiv.org/content/10.1101/2020.04.19.049452v2`.

65. Fong T, Jury B, Hu H, Murphy TH. PyRodentTracks: flexible computer vision and RFID based system for multiple rodent tracking and behavioral assessment; 2022. Available from: `https://www.biorxiv.org/content/10.1101/2022.01.23.477395v3`.

66. Szechtman H, Dvorkin-Gheva A, Gomez-Marin A. A virtual library for behavioral performance in standard conditions—rodent spontaneous activity in an open field during repeated testing and after treatment with drugs or brain lesions. GigaScience. 2022;11:giac092. doi:10.1093/gigascience/giac092.

67. Schweinfurth MK, Stieger B, Taborsky M. Experimental evidence for reciprocity in allogrooming among wild-type Norway rats. Scientific Reports. 2017;7(1):4010. doi:10.1038/s41598-017-03841-3.

68. Puścian A, Knapska E. Blueprints for measuring natural behavior. iScience. 2022;25(7). doi:10.1016/j.isci.2022.104635.

69. Grieco F, Bernstein BJ, Biemans B, Bikovski L, Burnett CJ, Cushman JD, et al. Measuring Behavior in the Home Cage: Study Design, Applications, Challenges, and Perspectives. Frontiers in Behavioral Neuroscience. 2021;15:735387. doi:10.3389/fnbeh.2021.735387.

70. Gunaydin LA, Grosenick L, Finkelstein JC, Kauvar IV, Fenno LE, Adhikari A, et al. Natural neural projection dynamics underlying social behavior. Cell. 2014;157(7):1535–1551.

71. Padilla-Coreano N, Tye KM, Zelikowsky M. Dynamic influences on the neural encoding of social valence. Nature Reviews Neuroscience. 2022;23(9):535–550.

72. Lee CR, Chen A, Tye KM. The neural circuitry of social homeostasis: Consequences of acute versus chronic social isolation. Cell. 2021;184(6):1500–1516.

73. Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nature Neuroscience. 2018;21(9):1281–1289. doi:10.1038/s41593-018-0209-y.

74. Graving JM, Chae D, Naik H, Li L, Koger B, Costelloe BR, et al. DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. eLife. 2019;8:e47994. doi:10.7554/eLife.47994.

75. Pereira TD, Aldarondo DE, Willmore L, Kislin M, Wang SSH, Murthy M, et al. Fast animal pose estimation using deep neural networks. Nature Methods. 2019;16(1):117–125. doi:10.1038/s41592-018-0234-5.

76. Pereira TD, Tabris N, Matsliah A, Turner DM, Li J, Ravindranath S, et al. SLEAP: A deep learning system for multi-animal pose tracking. Nature Methods. 2022;19(4):486–495. doi:10.1038/s41592-022-01426-1.

77. Lehner SR, Rutte C, Taborsky M. Rats Benefit from Winner and Loser Effects. Ethology. 2011;117(11):949–960. doi:10.1111/j.1439-0310.2011.01962.x.

78. Ramos A, Berton O, Mormède P, Chaouloff F. A multiple-test study of anxiety-related behaviours in six inbred rat strains. Behavioural brain research. 1997;85(1):57–69.

79. Neumann C. EloChoice: Preference Rating for Visual Stimuli Based on Elo Ratings; 2019. Available from: `https://cran.r-project.org/web/packages/EloChoice/index.html`.

80. Neumann C. EloSteepness: Bayesian Dominance Hierarchy Steepness via Elo Rating and David's Scores; 2023. Available from: `https://cran.r-project.org/web/packages/EloSteepness/index.html`.